



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Self-Reported Medication Use Validated Through Record Linkage to National Prescribing Data

Citation for published version:

Hafferty, J, Campbell, A, Navrady, L, Adams, MJ, Macintyre, D, Lawrie, S, Nicodemus, K, Porteous, DJ & McIntosh, A 2017, 'Self-Reported Medication Use Validated Through Record Linkage to National Prescribing Data', *Journal of Clinical Epidemiology*. <https://doi.org/10.1016/j.jclinepi.2017.10.013>

Digital Object Identifier (DOI):

[10.1016/j.jclinepi.2017.10.013](https://doi.org/10.1016/j.jclinepi.2017.10.013)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Journal of Clinical Epidemiology

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

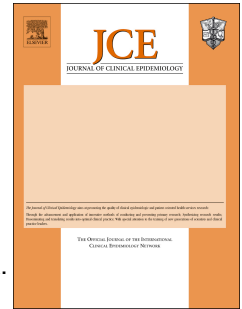
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Accepted Manuscript

Self-Reported Medication Use Validated Through Record Linkage to National Prescribing Data

J.D. Hafferty, A.I. Campbell, L.B. Navrady, M.J. Adams, D. MacIntyre, S.M. Lawrie, K. Nicodemus, D.J. Porteous, A.M. McIntosh



PII: S0895-4356(17)30306-2

DOI: [10.1016/j.jclinepi.2017.10.013](https://doi.org/10.1016/j.jclinepi.2017.10.013)

Reference: JCE 9509

To appear in: *Journal of Clinical Epidemiology*

Received Date: 24 March 2017

Revised Date: 10 October 2017

Accepted Date: 23 October 2017

Please cite this article as: Hafferty J, Campbell A, Navrady L, Adams M, MacIntyre D, Lawrie S, Nicodemus K, Porteous D, McIntosh A, Self-Reported Medication Use Validated Through Record Linkage to National Prescribing Data, *Journal of Clinical Epidemiology* (2017), doi: 10.1016/j.jclinepi.2017.10.013.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Self-Reported Medication Use Validated Through Record Linkage to National Prescribing Data

¹*Hafferty JD, ²Campbell AI, ¹Navrady LB, ¹Adams MJ, ¹MacIntyre D, ¹Lawrie SM, ^{2,3}Nicodemus K, ^{2,3,4}Porteous DJ, ^{1,4}McIntosh AM

1. Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh EH10 5HF
2. Generation Scotland, Centre for Genomics and Experimental Medicine, Institute for Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU
3. Institute for Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU
4. Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, 7 George Square, EH8 9JZ

*Corresponding author
Division of Psychiatry
University of Edinburgh
Royal Edinburgh Hospital
Edinburgh EH10 5HF

Corresponding author : jonathan.hafferty@ed.ac.uk
Tel : +44 (0)131 537 6260

Summary	200 words
Introduction	565 words
Methods	929 words
Results	584 words
Discussion	1649 words
TOTAL	3927 words
References	<u>47</u>

ACKNOWLEDGEMENTS

This investigation was supported by Wellcome Trust Grant 104036/Z/14/Z, and by the Dr Mortimer and Theresa Sackler Foundation. Generation Scotland received core support from the Chief Scientist Office of the Scottish Government Health Directorates [CZD/16/6] and the Scottish Funding Council [HR03006]. Dr Hafferty is supported by a Medical Research Council/Medical Research Foundation PsySTAR Fellowship.

Conflicts of Interest : None.

Summary

Objective

Researchers need to be confident about the reliability of epidemiological studies that quantify medication use through self-report. Some evidence suggests that psychiatric medications are systemically under-reported. Modern record linkage enables validation of self-report with national prescribing data as gold standard. Here, we investigated the validity of medication self-report for multiple medication types.

Study Design and Setting

Participants in the Generation Scotland population-based cohort (N=10,244) recruited 2009-11 self-reported regular usage of several commonly prescribed medication classes. This was matched against Scottish NHS prescriptions data using three- and six-month fixed time windows. Potential predictors of discordant self-report, including general intelligence and psychological distress, were studied via multivariate logistic regression.

Results

Antidepressants self-report showed very good agreement ($\kappa=0.85$, (95% Confidence Interval (CI) 0.84-0.87)), comparable to antihypertensives ($\kappa=0.90$, (0.89-0.91)). Self-report of mood stabilizers showed moderate-poor agreement ($\kappa=0.42$ CI 0.33-0.50). Relevant past medical history was the strongest predictor of self-report sensitivity, whereas general intelligence was not predictive.

Conclusion

In this large population-based study, we found self-report validity varied among medication classes, with no simple relationship between psychiatric medication and under-reporting. History of indicated illness predicted more accurate self-report, for both psychiatric and non-

psychiatric medications. Although other patient-level factors influenced self-report for some medications, none predicted greater accuracy across all medications studied.

Keywords

Agreement; Pharmacoepidemiology; Self-report; Medicines; Indication; Linkage

Word Count: 3927 words

WHAT IS NEW

- Self-reported medication use shows high validity in the general population although there is variation between medication classes.
- A simple relationship between psychiatric medications and under-reporting was not found. Mood stabilizers show moderate-poor agreement, due to both under-report and false positives, whereas antidepressant reporting is comparable to other long-term non-psychiatric medications.
- Medical history of an indicated health condition is the strongest predictor of accurate report. General intelligence was not associated with the accuracy of reporting.
- Medication-related factors such as range of indications, prescribing cycles, and phrasing of self-report question may also influence accuracy of self-report.
- When matching self-report to prescribing data, longer fixed time windows produce higher levels of agreement and positive predictive values, at the expense of some loss of sensitivity.

1 Introduction

Cohort studies, and other epidemiological studies using self-reported data, depend on the accuracy of the self-report to make accurate and reliable conclusions. This includes pharmaco-epidemiological and large-scale biobanking studies which are based on self-reported medication use. Self-reported medication use can be determined by questionnaire [1, 2]; by telephone or internet survey[3]; or by face-to-face interview[4-7]. However, self-report is subject to recall errors and biases[8, 9] and patients may be less willing to disclose details of certain medications than others.

The accuracy of self-report can be verified by comparison to a trusted measure or “gold standard”. For medication utilization, the choice of gold standard depends to an extent on the purpose of the study (i.e. estimating patient adherence, or monitoring prescribing behaviour of clinicians), and there is therefore no universally applicable and accepted gold standard [10] [11]. One option is for a third party to perform a home inventory [12] or record individual medications produced by the patient [13], but these assessments are difficult to perform on a large scale. An alternative is to compare self-report data to prescriptions, healthcare insurance claims, or general practice medical records [4, 5, 11, 14]. Prescribing databases have been shown to be highly accurate in recording medication utilization [15], at least for those medications that require prescriptions.

Among published studies comparing medication self-report to prescribing data, the majority have been relatively small in size[4, 6, 7, 10-13, 16-18]. Many studies are restricted to certain medications or medication types, such as antihypertensives [11]; cardiovascular drugs [6]; antidepressants [17], or hormone replacement therapy [1]; or to special populations, such as the elderly [6, 12, 15]; postmenopausal women [2, 5]; or psychiatric illnesses [16]. Few

studies utilize large population-based samples[4, 13, 14, 19] or multiple disparate medication types[13, 19-21]. Such comparisons are important, however, for they enable study of systematic over- and under-reporting of medication utilization between drug classes.

Self-report can be compromised by a number of factors, including not understanding the question, poor recall, and intended non-disclosure[4]. There is no consensus on patient-level factors predisposing to discordance between medication self-report and gold standard measures, but previous reports have implicated advancing age[9, 19], being unmarried [19, 21], number of medications regularly dispensed[18, 22], suffering poor health [19], and lower educational attainment[21]. Within medication classes, there is some evidence that psychiatric medications are less likely to be accurately self-reported[19, 22]. Potential explanations for this include confusion regarding medication indication but also non-disclosure due to social desirability bias[9] or self-stigmatization[2, 4, 10, 23]. Factors that have not to date been found to influence reporting include gender[19, 21] and cognitive health [21].

Prescribing data can be sourced from local health providers or insurers[10], pharmacy records [6, 11, 13, 14, 17, 21], social insurance databases [16, 19] or national health service databases [1, 2, 4]. The recording of the dispensing and collection of medication, as well as its prescribing, is important for studies that seek to measure patient utilization (although even collection of a medication is not a hard indicator of usage). The country of origin of the study, and respective prescription legislation, dispensing and reimbursement practices, are also relevant to interpreting self-report against prescribing data (for example, over-the-counter medications may not appear in this data), and to making comparisons between national studies.

In this study, we sought to ascertain agreement between medication self-report, derived from a large UK cohort study, compared to record-linked national prescribing data as gold standard, across a range of commonly used psychiatric and non-psychiatric medications. We hypothesised that agreement would be lower for psychiatric medication types, due to systemic under-reporting. To our knowledge this is one of the largest population-based studies of medication self-report also incorporating a covariate analysis method across a range of medications.

2 Method

2.1 Study Population

Our study utilized the Generation Scotland: Scottish Family Health Study (GS:SFHS) family- and population-based cohort of Scottish adult volunteers (n=21,474), recruited February 2006-March 2011, which has been described elsewhere[24, 25]. The cohort has a higher proportion of females (59%) and older median age (47 males: 48 females) than the Scottish population at the 2001 census (37 and 39 respectively) [25, 26]. Written informed consent was obtained for 98% of GS:SFHS for data linkage to routinely collected healthcare records.

2.2 Medication Self Report Data

All participants in GS:SFHS were asked to complete a pre-clinic questionnaire prior to their enrolment in the study. The first phase of the study used a text-based questionnaire which is not part of this analysis. Those individuals recruited between June 2009 – March 2011 (n=10,980, 59.5% female) completed a coded questionnaire where the Medications section was a “Yes” versus “No” checkbox, with the accompanying question “*Are you regularly taking any of the following medications?*”. The available options were: (1) “*Cholesterol lowering medication (e.g. Simvastatin)*” (2) “*Blood pressure lowering medication*” (3) “*Insulin*” (4) “*Hormone replacement therapy*” (5) “*Oral contraceptive pill or mini pill*” (5) “*Aspirin*” (6) “*Antidepressants*” (7) “*Mood stabilizers*”. The completed questionnaires were then machine read and electronically recorded using anonymised patient linkers.

2.3 Additional Covariate Data

Additional sociodemographic information collected in the questionnaire included gender, age, educational attainment, smoking status and relationship status. Compared to the rest of GS, our sample was moderately older and contained more individuals with no school qualifications and also more degree level educated individuals (Table 1, Figure 1). Lifetime history of affective disorder (major depression and bipolar disorder) was obtained using the Structured Clinical Interview for DSM-IV Disorders (SCID)[25]. Self-reported history of hypertension, heart disease and diabetes was recorded. In addition, during the GS interview a variety of cognitive tests were performed [24], including digit symbol from the Wechsler Adult Intelligence Scale III [27], logical memory from the Wechsler Memory Scale III[28] and verbal fluency[29]. From these tests, we derived a measure of general intelligence (g) as the first un-rotated principle component, explaining 44% of the variance in scores. [30, 31] Psychological distress was measured using the General Health Questionnaire-28(Likert scoring)[32].

2.4 Prescribing Data and Linkage

All Scottish citizens registered with a General Practitioner (more than 96% of the population) are assigned a unique identifier (Community Health Index (CHI) number). This was employed to record link GS:SFHS questionnaire data to the national Prescribing Information System (PIS) administered by NHS Services Scotland Information Services Division [33]. PIS is a database of all Scottish NHS prescriptions for payments for medications prescribed by GPs; nurses; dentists; pharmacists; and hospitals where the medication was dispensed in the community. There is no prescription charge in Scotland. Hospital dispensed prescriptions and over-the-counter medications are not included. Patient level data has been available in

PIS since April 2009 [34]. We obtained PIS prescribing data for April 2009-March 2011. We used the dates of dispensing, not prescription, when matching to self-report.

2.5 Matching Prescribing to Self-Report

For each individual and medication type, concordance with GS:SFHS self-report was checked against PIS prescribing record dispensing dates within a “fixed time window” [2, 4, 14, 16] including the month of questionnaire completion, and two months preceding (total three months), and also five months preceding (total six months). The majority of prescriptions, including in Scotland, are dispensed in quantities of 90 days duration or less [13, 35]. A previous Dutch study [12] also found that fixed time windows shorter than 90 days are less sensitive, although the generalizability of this finding is uncertain. Accordingly, we employed two fixed time windows, three and six months duration, in order to assess their relative benefits in terms of agreement, sensitivity and positive predictive value.

To ensure all individuals had at least six months of potentially available prescribing records, we restricted analysis to GS:SFHS participants who had completed their medication questionnaire in September 2009 or later. This equated to 10, 244 participants (6065 females and 4179 males) enrolled September 2009-March 2011 (Table 1, Figure 1). Of these, 96.5% had medication records available (the remainder were presumably not using prescribed medication) which compared to 95.6% for the whole GS cohort.

The PIS data allows medications to be identified by approved drug name and/or associated British National Formulary [36] paragraph code. Medication indication is not recorded. Our matching criterion for each medication type is detailed in Table 4 (Supplementary).

2.6 Missing Data

The self-report questionnaire employed a ‘Yes’/‘No’ checkbox, but some individuals ticked neither box (or data was otherwise missing, Table 2). In our main analysis we treated each medication separately, excluding the missing self-report values for each case. However, to mitigate the potential of hereby introducing biases, or not accounting for individuals who intended to deny medication use by leaving the section blank, we conducted two additional analyses – one with all individuals with any missing data excluded ($n=7836$), and the other with missing data coded as denial of medication use (Table 5, Supplementary).

2.7 Statistical Analysis

All analyses were carried out using R version 3.2.3[37]. Level of agreement between self-report and prescribing data was ascertained using Cohen’s kappa (κ) method of rating inter-observer variation [38]. Kappa scores of <0.40 were considered fair to poor; $0.41-0.60$ moderate; $0.61-0.8$ substantial; and >0.81 good or better [39, 40]. We also calculated sensitivity, specificity and positive predictive values (PPV). Ninety five percent confidence intervals (CI) were included.

We performed multivariate logistic regression analysis on predictors of false negative self-report compared to true positive (sensitivity). Due to some covariate missing data, the sample size of this analysis was reduced to 9043 for complete case analysis (Table 1, Figure 1). Odds ratios with 95% CI were calculated. Multiple testing was adjusted for using the False Discovery Rate method with significance level (α) 0.05. As Generation Scotland is a

partly family-based cohort, we adjusted for any correlation due to family relatedness using the Generalized Estimating Equations method[41].

3 Results

Of the 10,244 individuals in the study, 6164 (60.17%) ticked 'No' to every medication question (Figure 1). In addition, 485 (4.74%) left blank or had missing data for every question. The proportion of completed responses differed between medications and was greatest for antihypertensives (86.44%) and lowest for mood stabilizers (77.87%, $\chi^2 = 256.07$, $p < 2.2E-16$) (Table 2). The most commonly prescribed medication (six-month window) was antihypertensives, prevalence 19.05%, whereas antidepressants prevalence was 12.22% and mood stabilizers 1.32%. The prevalence of lifetime history of affective disorder in our sample was 12.66% ($n=1297$) for major depressive disorder and 0.31% for bipolar disorder ($n=32$). The self-reported prevalence of hypertension was 12.66% ($n=1297$), heart disease 3.37% ($n=345$) and diabetes 3.15% ($n=323$) (Table 1).

3.1 Agreement and Validity

Agreement (Table 2, Figure 2) between medication self-report and prescribing data was generally very good across medication classes. Greatest agreement was found for cholesterol lowering medication ($\kappa=0.95$, CI 0.94-0.96) (6-month fixed time window unless otherwise stated). Agreement for antidepressants ($\kappa=0.85$, CI 0.84-0.87) was lower than antihypertensives ($\kappa=0.90$, CI 0.89-0.91) but still within the highest kappa banding of >0.81 . By contrast, agreement for mood stabilizers was moderate-poor ($\kappa=0.42$, CI 0.33-0.50). Comparing the six-month fixed time window to three-month, κ scores were higher, although only to a degree beyond 95% confidence intervals in the case of HRT and oral contraceptives.

Self-report sensitivity (Table 2, Figure 2) was slightly reduced in the six-month time window versus three-month, but was still greater than 0.80 for all medications except mood stabilizers. Antidepressant sensitivity (0.85, CI 0.82-0.87) was comparable to antihypertensives (0.86, CI 0.85-0.88). Sensitivity for mood stabilizers was comparatively poor (0.40, CI 0.31-0.50) indicating a high rate of false negatives.

The positive predictive value (Table 2, Figure 3) for antidepressant use (0.89, CI 0.87-0.91) was substantial, albeit less than antihypertensives and cholesterol lowering drugs, and contrasted with modest PPV for mood stabilizers (0.45 CI 0.35-0.56). The six-month fixed time window significantly improved PPV for most medication groups, with greatest effect for HRT and oral contraceptives (which nevertheless showed relatively moderate PPV in both time windows).

3.2 Predictors of Failure To Self-Report Medication Usage

Multivariate logistic regression (Table 3) found no covariates universally associated, across all medications, with failure to self-report medication usage, as determined by the prescribing data gold standard. General intelligence (g) was not associated with increased false negatives for any medication. Psychological distress (GHQ) reduced odds of false negatives for antidepressants (OR 0.98, CI 0.96-1.00, p_{FDR} 0.081) and mood stabilizers (OR 0.96 CI 0.91-1.01, p_{FDR} 0.197), but this relationship was not significant for multiple testing.

There was reduced discordant self-reporting for several medications if the patient had a history of an illness for which that medication was indicated, such as affective disorder and mood stabilisers (OR 0.09, CI 0.02-0.35 p_{FDR} 0.005), and hypertension and antihypertensives

(OR 0.04, CI 0.02-0.06 $p_{FDR} > 0.001$). Similar associations were found for affective disorder and antidepressants, and cardiac disease and aspirin, with p values of > 0.1 after correcting for multiple testing.

Age and gender showed no consistent association, although older age was associated with lower false negatives for antihypertensives, antidepressants and possibly aspirin (p_{FDR} 0.074), and female gender was associated with increased false negatives for antihypertensives (OR 1.75, CI 1.16-2.62, p_{FDR} 0.020).

3.3 Influence of Missing Data

Recoding missing data as negative self-report (Table 5, Supplementary) resulted in somewhat lower levels of agreement and lower sensitivities for all medications. However, agreement remained good for antidepressants ($\kappa = 0.81$ CI 0.79-0.83) and poor for mood stabilisers (0.34 CI 0.26-0.41). There was a demonstrable reduction in sensitivity for antidepressants (0.78 CI 0.75-0.80) but this reduction was not confined to psychiatric medications, being found also in antihypertensives (0.79 CI 0.77-0.81).

4 Discussion

In this population-based cohort, we found substantial to very good agreement between medication self-report and electronic prescribing records, for most medications studied. We hypothesised that psychiatric medications would show less agreement and systematic under-reporting. Agreement for mood stabilizers was indeed considerably worse, although we found evidence of both under- and over-reporting(false positives). However, for antidepressants the agreement, sensitivity and PPV were broadly comparable to other medications studied. We did not identify any generalizable single predictors of failure to self-report prescribed medications, for psychiatric medications or for medications generally. However, past medical history of an indicated health condition showed the strongest effect in promoting self-report accuracy across classes, and this was also true for psychiatric medications.

In general, the six-month fixed time window outperformed the three-month for agreement and PPV, at the expense of modest loss of sensitivity. This was most evident for HRT and oral contraceptives in women, which could imply these medications are dispensed in longer time cycles, and require longer fixed time windows relative to other medications.

4.1 Predictors of Discordant Self-Report

We found that a medical history of an indicated health condition for a given medication, such as affective disorder for mood stabilizers, or hypertension for antihypertensives, reduced the odds of false negatives. If systematic under-reporting of psychiatric medications due to self-stigma was taking place, we might have expected to find the reverse. Relationship status and educational status did not predict discordance, except in the case of mood stabilizers where lack of school qualifications was associated with false negative reporting. This could indicate

reduced understanding of the definition of “mood stabilizer” among the less educated, but might also represent association between lesser educational achievement and use of medications (such as antipsychotics) included in our definition of mood stabilizers.

We found that general intelligence (g) did not influence concordance of medication self-report with prescribing data, which to our knowledge has not been previously reported. We also believe we are the first to investigate psychological distress and medication self-report. Interestingly, while psychological distress might be posited as a potential factor in under-reporting psychiatric medications (e.g. through self-stigma), we found some evidence of a relationship between the increased GHQ score and greater sensitivity of self-reporting of antidepressants ($p < 0.1$). Gender was not generally associated with accuracy, except in the case of antihypertensives, where increased odds of false negatives (OR 1.75 CI 1.16-2.62) were found, perhaps indicating greater usage of these medication types for non-antihypertensive purposes among females.

4.3 Questionnaire Phrasing

One possible explanation for the poor agreement, sensitivity and PPV for mood stabilizers is confusion among questionnaire respondents about the meaning of “mood stabilizer”. There is no consensus definition of mood stabilizer among clinicians,[42] and laypersons may therefore be unsure as to its meaning. Klungel[8] has previously reported that sensitivity of medication self-report is influenced by the specificity of question phrasing. In our matching to prescribing data we employed a broad definition of mood stabilizers, but when a narrower definition (excluding antipsychotics) was employed the agreement was even worse ($\kappa = 0.29$, CI 0.20-0.38).

4.4 Comparison With Other Studies

Table 6(Supplementary) describes the agreement of this present study, using the 6-month fixed time window, with other large published studies. We report a higher level of agreement ($\kappa=0.86$) for antidepressants than Nielsen ($\kappa=0.66$)[4], Rauma ($\kappa=0.65$) [2] and Noize ($\kappa=0.81$)[20]. When making comparisons with studies performed in other healthcare systems, it is important recognise the variations between countries in prescribing legislation and access to medication. Scotland has a national health system, with no prescription charges, and prescribing data is collated nationally, which might explain a higher concordance with self-report and prescribing data than might be possible in some comparator studies.

Kwon[10] compared survey antidepressant self-report in a longitudinal depression study ($n=164$) with pharmacy claims data and a three-month fixed window and found substantial levels of agreement ($\kappa=0.69$). Interestingly, where there were discrepancies in prescription record antidepressant use, they found on notes review that most cases could be explained by antidepressants being used for other indications, or due to recent discontinuation. In our study, we attempted to minimise the rate of antidepressant false positives due to other indications by excluding amitriptyline from our searches (amitriptyline is widely prescribed but now rarely for depression in the UK).

With regard to mood stabilizers, a recent study comparing self-reported medication use in a genetic study of schizophrenia ($n=905$) [16] found substantial levels of agreement ($\kappa=0.74$) between self-report of mood stabilizers and an administrative prescription database. This is a much higher level of agreement than found in our study, although we note that Haukka's was not a community-based sample and had a much higher prevalence of mood stabilizer use. A comparison of a postal medication survey ($n=11,031$) with national prescription records

reported by Rauma[2] found substantial levels of agreement for antidepressant reporting ($\kappa=0.65$) but poor agreement ($\kappa=0.30$) for other psychoactive medications, a result more comparable with our own findings.

4.5 Study Strengths and Weaknesses

Our study used a large ($n=10,244$) population-based cohort linked to high fidelity Scottish PIS records (capture rate in excess of 95%)[34]. Self-report was via a short, simply worded questionnaire which obviated interviewer bias and did not require long-term recall of medication use. Response rate was high. We employed a variety of methods to compare the two data sources over two fixed time windows and performed covariate analysis of predictors of discordant self-report.

However, our method of verifying medication utilization took no account of dose and concordance with medication was assumed. Patients may be prescribed a drug but not fill their prescription (primary noncompliance), although our use of date of dispensing rather than prescribing date would have obviated this to an extent, it would still be unknown if the dispensed drug was collected. In addition, patients may not take the drug, or not take as intended (secondary noncompliance), and concordance can be as low as 50% for antidepressants and antihypertensives[4, 43]. In addition, the questionnaire referred to “regularly” taken medication whereas our method recorded any prescription within the fixed time window as positive use. The absence of data in PIS on medication indication increased the risk of over-inclusion and false positives, particularly for medications with broader indications, although we attempted to decrease this using our exclusion criteria (Table 4, Supplementary). Fixed time windows also potentially record false positives for medications

discontinued during the window, but prior to self-report, although this is more common with medications taken acutely, such as antibiotics[12].

We must therefore concede that prescription data is by its nature an imperfect gold standard, although its use enables very large sample sizes which improve overall accuracy. The use of prescribing data as a gold standard involves some strong assumptions, including that the patient could not have obtained the medication without it being recorded in the prescribing data. The extent to which this is true depends on a variety of variables, including the medication type, prescribing legislation of the country of study, and the movement of individual patients between healthcare providers. Indeed, some studies are performed on the basis of self-report as gold standard to analyse the validity of clinical or prescribing records[44]. However, the advantage of prescribing data as a gold standard is that it is an objective measure, with definitions of medication usage that can be readily replicated across studies and countries (whereas self-report questionnaires can vary considerably in definition and interpretation), which can be utilised at large scale across multiple medication types, and that is not subject to potential recall and desirability biases of self-report studies[45].

Data linkage is also a fast-moving field, and though the PIS data from 2011 we used in this study had high fidelity and a capture in excess of 95%, future studies using larger datasets and more complex linkage may enable even more accurate estimates of validity. For example, as data linkage improves, cross referencing to other sources of clinical data such as GP and hospital records should assist identifying true cases and also reduce the incidence of false positives for those who have discontinued medication through the time windows analysed.

As discussed, the use of the term “mood stabilizer” may have caused confusion. Many individuals did not tick either checkbox, and response rate differed between medication types, from 86.44% for antihypertensives to 77.87% for mood stabilizers. This may have reflected variations in understanding of, or willingness to answer, the question, and could have biased our results or inflated the kappa scores. However, we demonstrated that recoding this missing data as denial of use still produced substantial levels of agreement (Table 5, Supplementary). The Cohen’s kappa method itself may inflate values depending on the proportion of subjects in each category[46], hence we have also tabulated the raw proportions (Table 7, Supplementary). GS:SFHS is a partly family-based cohort and this could potentially have introduced some correlation bias into our analysis, although we accounted for this in our multivariate regression through Generalized Estimating Equations.

4.6 Conclusion

Our study provides convincing evidence that medication self-report is accurate compared to prescribing data, particularly for medication classes that are more precisely definable. We have shown that self-report of antidepressant use meets the highest threshold for Cohen’s kappa agreement and can be considered valid for research and clinical purposes. Our analysis of potential patient-level predictors of reporting discordance, such as gender, age, education and general intelligence, did not identify generalizable factors across all medication classes, although there was some evidence that medical history of an indicated condition improves sensitivity of self-report. As discussed above, medication-level factors such as range of possible indications, and length of dispensing cycles, may also be important when validating self-report across a fixed time window with prescribing data as gold standard.

Our study also demonstrates the utility of record linkage of longitudinal population based cohorts to nationally administered prescribing datasets, as a useful adjunct to epidemiological and large biobanking studies. Utilising administrative health data for verification and quality control of self-report has applications beyond epidemiological studies and can be potentially exploited in clinical applications, such as data-linked clinical support tools acting as adjuncts to clinical interview, and in formulating predictive models of disease risk.[47]

References

- [1] Lokkegaard EL, Johnsen SP, Heitmann BL, Stahlberg C, Pedersen AT, Obel EB, et al. The validity of self-reported use of hormone replacement therapy among Danish nurses. *Acta Obstet Gynecol Scand.* 2004;83:476-81.
- [2] Rauma PH, Koivumaa-Honkanen H, Kroger H, Tuppurainen MT, Kauhanen J, Honkanen RJ. The relationship between self-reported and registry-based data on use of psychoactive medications in postmenopausal women. *Bmc Psychiatry.* 2013;13.
- [3] West SL, Savitz DA, Koch G, Strom BL, Guess HA, Hartzema A. Recall accuracy for prescription medications: self-report compared with database information. *Am J Epidemiol.* 1995;142:1103-12.
- [4] Nielsen MW, Sondergaard B, Kjoller M, Hansen EH. Agreement between self-reported data on medicine use and prescription records vary according to method of analysis and therapeutic group. *J Clin Epidemiol.* 2008;61:919-24.
- [5] Goodman MT, Nomura AM, Wilkens LR, Kolonel LN. Agreement between interview information and physician records on history of menopausal estrogen use. *Am J Epidemiol.* 1990;131:815-25.
- [6] Sjahid SI, van der Linden PD, Stricker BH. Agreement between the pharmacy medication history and patient interview for cardiovascular drugs: the Rotterdam elderly study. *Br J Clin Pharmacol.* 1998;45:591-5.
- [7] Norell SE, Boethius G, Persson I. Oral contraceptive use: interview data versus pharmacy records. *Int J Epidemiol.* 1998;27:1033-7.
- [8] Klungel OH, de Boer A, Paes AH, Herings RM, Seidell JC, Bakker A. Influence of question structure on the recall of self-reported drug use. *J Clin Epidemiol.* 2000;53:273-7.
- [9] Cotterchio M, Kreiger N, Darlington G, Steingart A. Comparison of self-reported and physician-reported antidepressant medication use. *Ann Epidemiol.* 1999;9:283-9.
- [10] Kwon A BK, Pei Y, Rogers WH, Wilson IB, Zhou Q, MD, Adler DA. Antidepressant Use Concordance Between Self-Report and Claims Records. *MEDICAL CARE.* 2003;41:368-74.
- [11] Klungel OH, de Boer A, Paes AH, Herings RM, Seidell JC, Bakker A. Agreement between self-reported antihypertensive drug use and pharmacy records in a population-based study in The Netherlands. *Pharm World Sci.* 1999;21:217-20.
- [12] Lau HS, de Boer A, Beuning KS, Porsius A. Validation of pharmacy records in drug exposure assessment. *J Clin Epidemiol.* 1997;50:619-25.
- [13] Caskie GI, Willis SL, Warner Schaie K, Zanjani FA. Congruence of medication information from a brown bag data collection and pharmacy records: findings from the Seattle longitudinal study. *Exp Aging Res.* 2006;32:79-103.
- [14] Monster TB, Janssen WM, de Jong PE, de Jong-van den Berg LT, RENal PSGPo, Vascular ENTSD. Pharmacy data in epidemiological studies: an easy to obtain and reliable tool. *Pharmacoepidemiol Drug Saf.* 2002;11:379-84.
- [15] Tamblyn R, Lavoie G, Petrella L, Monette J. The use of prescription claims databases in pharmacoepidemiological research: the accuracy and comprehensiveness of the prescription claims database in Quebec. *J Clin Epidemiol.* 1995;48:999-1009.
- [16] Haukka J, Suvisaari J, Tuulio-Henriksson A, Lonnqvist J. High concordance between self-reported medication and official prescription database information. *Eur J Clin Pharmacol.* 2007;63:1069-74.

- [17] Saunders K, Simon G, Bush T, Grothaus L. Assessing the feasibility of using computerized pharmacy refill data to monitor antidepressant treatment on a population basis: a comparison of automated and self-report data. *J Clin Epidemiol*. 1998;51:883-90.
- [18] Jain MG, Rohan TE, Howe GR. Agreement of self reported use of menopausal hormone replacement therapy with physician reports. *Epidemiology*. 1999;10:260-3.
- [19] Haapea M, Miettunen J, Lindeman S, Joukamaa M, Koponen H. Agreement between self-reported and pharmacy data on medication use in the Northern Finland 1966 Birth Cohort. *Int J Methods Psychiatr Res*. 2010;19:88-96.
- [20] Noize P, Bazin F, Dufouil C, Lechevallier-Michel N, Ancelin ML, Dartigues JF, et al. Comparison of health insurance claims and patient interviews in assessing drug use: data from the Three-City (3C) Study. *Pharmacoepidemiol Drug Saf*. 2009;18:310-9.
- [21] Richardson K, Kenny RA, Peklar J, Bennett K. Agreement between patient interview data on prescription medication use and pharmacy records in those aged older than 50 years varied by therapeutic group and reporting of indicated health conditions. *Journal of clinical epidemiology*. 2013;66:1308-16.
- [22] Van den Brandt PA, Petri H, Dorant E, Goldbohm RA, Van de Crommert S. Comparison of questionnaire information and pharmacy data on drug use. *Pharm Weekbl Sci*. 1991;13:91-6.
- [23] Knudsen P. Perceptions of young women using SSRI antidepressants - a reclassification of stigma. *Int J Pharm Pract*. 2002;10:243-52.
- [24] Smith BH, Campbell H, Blackwood D, Connell J, Connor M, Deary IJ, et al. Generation Scotland: the Scottish Family Health Study; a new resource for researching genes and heritability. *BMC Med Genet*. 2006;7:74.
- [25] Smith BH, Campbell A, Linksted P, Fitzpatrick B, Jackson C, Kerr SM, et al. Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. *Int J Epidemiol*. 2013;42:689-700.
- [26] Smith BH CA, Linksted P, Fitzpatrick B, Jackson C, Kerr SM, Deary IJ, MacIntyre DJ, Campbell H, McGilchrist M, Hocking LJ, Wisely L, Ford I, Lindsay RS, Morton R, Palmer CNA, Dominiczak AF, Porteous DJ, Morris AD. Cohort profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. *International Journal of Epidemiology*. 2012:1-12.
- [27] Wechsler D. Wechsler Memory Scale III, . London: Psychological Corporation; 1998.
- [28] Wechsler D. Wechsler Adult Intelligence Scale III. London: Psychological Corporation; 1998.
- [29] Lezak MD. Neuropsychological Testing. 3rd ed. Oxford: Oxford University Press; 1995.
- [30] Navrady LB. Intelligence and neuroticism in relation to depression and psychological distress: evidence of interaction using data from Generation Scotland: Scottish Family Health Study and UK Biobank 2017.
- [31] Marioni RE, Batty GD, Hayward C, Kerr SM, Campbell A, Hocking LJ, et al. Common genetic variants explain the majority of the correlation between height and intelligence: the generation Scotland study. *Behav Genet*. 2014;44:91-6.
- [32] Goldberg DP, Hillier VF. A scaled version of the General Health Questionnaire. *Psychol Med*. 1979;9:139-45.
- [33] Alvarez-Madrado S, McTaggart S, Nangle C, Nicholson E, Bennie M. Data Resource Profile: The Scottish National Prescribing Information System (PIS). *Int J Epidemiol*. 2016.

- [34] Medicines for Mental Health : Financial Years 2004/05 to 2013/14: Information Services Division; 2014.
- [35] Reid I. Reponse to Spence D "Unhappy Pills". BMJ. 2012;345.
- [36] British National Formulary: Pharmaceutical Press; 2012.
- [37] R_Core_Team. R: A language and environment for statistical computing. Vienna, Austria. : R Foundation for Statisical Computing. ; 2015.
- [38] Cohen J. A Coefficient of Agreement for Nominal Scales. Educ Psychol Meas. 1960;20:37-46.
- [39] Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med. 2005;37:360-3.
- [40] Landis JR KG. The Measurement of Observer Agreement for Categorical Data. Biometrics 1977:159-74.
- [41] Hanley JA, Negassa A, Edwardes MD, Forrester JE. Statistical analysis of correlated data using generalized estimating equations: an orientation. Am J Epidemiol. 2003;157:364-75.
- [42] Bauer MS, Mitchner L. What is a "mood stabilizer"? An evidence-based response. Am J Psychiatry. 2004;161:3-18.
- [43] Haynes RB, Ackloo E, Sahota N, McDonald HP, Yao X. Interventions for enhancing medication adherence. Cochrane Db Syst Rev. 2008.
- [44] Rikala M, Hartikainen S, Sulkava R, Korhonen MJ. Validity of the Finnish Prescription Register for measuring psychotropic drug exposures among elderly finns: a population-based intervention study. Drugs Aging. 2010;27:337-49.
- [45] Lam WY, Fresco P. Medication Adherence Measures: An Overview. Biomed Res Int. 2015;2015:217047.
- [46] Thompson WD, Walter SD. A reappraisal of the kappa coefficient. J Clin Epidemiol. 1988;41:949-58.
- [47] McIntosh AM, Stewart R, John A, Smith DJ, Davis K, Sudlow C, et al. Data science for mental health: a UK perspective on a global challenge. Lancet Psychiatry. 2016;3:993-8.

Table 1. Socio-demographic, clinical and cognitive characteristics of study populations compared to whole Generation Scotland cohort.

	GS:SFHS (N=21474)	Individuals in the current study (N=10244)	Subset of individuals in current study used in complete case multivariate logistic regression analysis (N=9043)
Female	12674 (59.02%)	6065 (59.21%)	5329 (58.9%)
Age 18-39	6769 (31.52%)	3072(29.99%) †	2797(30.93%) ‡
Age 40-64	12346 (57.49%)	6015 (58.72%) †	5304(58.65%)
Age 65-99	2359 (10.99%)	1157(11.29%)	942(10.42%)
Affective Disorder (SCID)	2848 (13.26%)	1329 (12.97%)	1159 (12.82%)
Diabetes (Self-Report)	659 (3.07%)	323 (3.15%)	277 (3.06%)
Hypertension(Self-Report)	2836 (13.21%)	1297 (12.66%) †	1125 (12.44%)
Cardiac Disease(Self-Report)	777 (3.62%)	345 (3.37%) †	284 (3.14%) ‡
No School Certificate	2452 (11.42%)	1432 (13.98%) †	1296 (14.33%) ‡
Postgraduate Education	6323 (29.44%)	3273 (31.95%) †	3164 (34.99%) ‡
Smoker	3662 (17.05%)	1733 (16.92%)	1484 (16.41%) ‡
Relationship Status - Single	6720 (31.29%)	3236 (31.59%)	2866 (31.69%) ‡
GHQ Likert Score	16 (8.87)	15.73 (8.74) †	15.66 (8.69) ‡
Wechsler Logical Memory Test I & II	30.7 (8.48)	30.95(8.15)	31.17 (8.05) ‡
Mill-Hill Vocabulary Test	30.06 (4.76)	30.09 (4.66)	30.23 (4.62) ‡
Wechsler Digit Symbol Substitution Task	72.23 (17.22)	71.71 (17.15) †	72.52 (16.88) ‡
Verbal Fluency Test	39.71 (11.72)	39.89 (11.70) †	40.22 (11.65) ‡

Abbreviations: GS:SFHS, Generation Scotland : Scottish Family Health Study. GHQ, General Health Questionnaire.

All values are totals with percentages, unless shown in italics where they are means with standard deviations in parentheses.

† = Significant differences (alpha=0.05) between Generation Scotland and Study Population as determined by Chi square / t tests.

‡ = Significant differences (alpha=0.05) between Study Population and subset used in multivariate logistic regression analysis as determined by Chi square / t tests.

Table 2. Medication self-report and prescribing data prevalences, agreements, sensitivities, specificities and positive predictive values, measured on two fixed time windows – 3 months and 6 months duration respectively – in the current study (n=10,244, including 6065 females)

	Total (n) completed question, with Yes or No (%)	Medication prevalence according to self report (%)	Medication prevalence according to PIS (%)*	3 MONTH FIXED TIME WINDOW				6 MONTH FIXED TIME WINDOW			
				Agreement κ (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Agreement κ (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)
Antidepressant**	8333 (81.35)	9.60	10.10	0.84 (0.82-0.86)	0.90 (0.87 – 0.92)	0.99 (0.99-0.99)	0.90 (0.87-0.92)	0.85 (0.84-0.87)	0.85 (0.82-0.87)	0.99 (0.99-0.99)	0.89 (0.87-0.91)
Mood stabilizer ***	7977 (77.87)	1.17	1.32	0.40 (0.31-0.49)	0.41 (0.31-0.52)	0.99 (0.99-0.99)	0.41 (0.31-0.52)	0.42 (0.33-0.50)	0.40 (0.31-0.50)	0.99 (0.99-1.00)	0.45 (0.35-0.56)
Cholesterol lowering medication	8789 (85.80)	13.97	13.81	0.92 (0.91-0.94)	0.97 (0.96-0.98)	0.98 (0.98-0.99)	0.90 (0.88-0.92)	0.95 (0.94-0.96)	0.97 (0.95-0.97)	0.99 (0.99-0.99)	0.95 (0.94-0.97)
Antihypertensive	8855 (86.44)	16.85	19.05	0.90 (0.89-0.91)	0.89 (0.87-0.91)	0.99 (0.99-0.99)	0.95 (0.94-0.96)	0.90 (0.89-0.91)	0.86 (0.85-0.88)	1.00 (0.99-1.00)	0.98 (0.97-0.98)
Aspirin	8445 (82.44)	9.28	7.63	0.81 (0.78-0.83)	0.97 (0.95-0.98)	0.97 (0.97-0.98)	0.72 (0.68-0.75)	0.84 (0.82-0.86)	0.95 (0.93-0.96)	0.98 (0.97-0.98)	0.78 (0.75-0.81)
Insulin	8016 (78.25)	1.11	0.97	0.87 (0.82-0.93)	1.00 (0.92-1.00)	1.00 (1.00-1.00)	0.78 (0.67-0.86)	0.93 (0.89-0.97)	1.00 (0.93-1.00)	1.00 (1.00-1.00)	0.88 (0.79-0.94)
HRT (female only)	*4794 (79.04)	5.97	4.59	0.62 (0.57-0.68)	0.92 (0.87-0.96)	0.97 (0.96-0.97)	0.49 (0.43-0.55)	0.78 (0.74-0.82)	0.91 (0.86-0.94)	0.98 (0.98-0.98)	0.70 (0.64-0.75)
Oral contraceptives (female only)	*4849 (79.95)	14.62	12.79	0.55 (0.51-0.59)	0.82 (0.78-0.86)	0.92 (0.91-0.92)	0.47 (0.43-0.51)	0.73 (0.70-0.76)	0.82 (0.79-0.85)	0.95 (0.95-0.96)	0.72 (0.68-0.75)

Abbreviations : PIS, Prescribing Information System. HRT, Hormone Replacement Therapy

* Six month time window employed

** Note that a broader definition of antidepressant than that shown in table, which included amitriptyline, returned an agreement of $\kappa=0.83(0.81-0.85)$ at six month time window with sensitivity of 0.75(0.73-0.78)

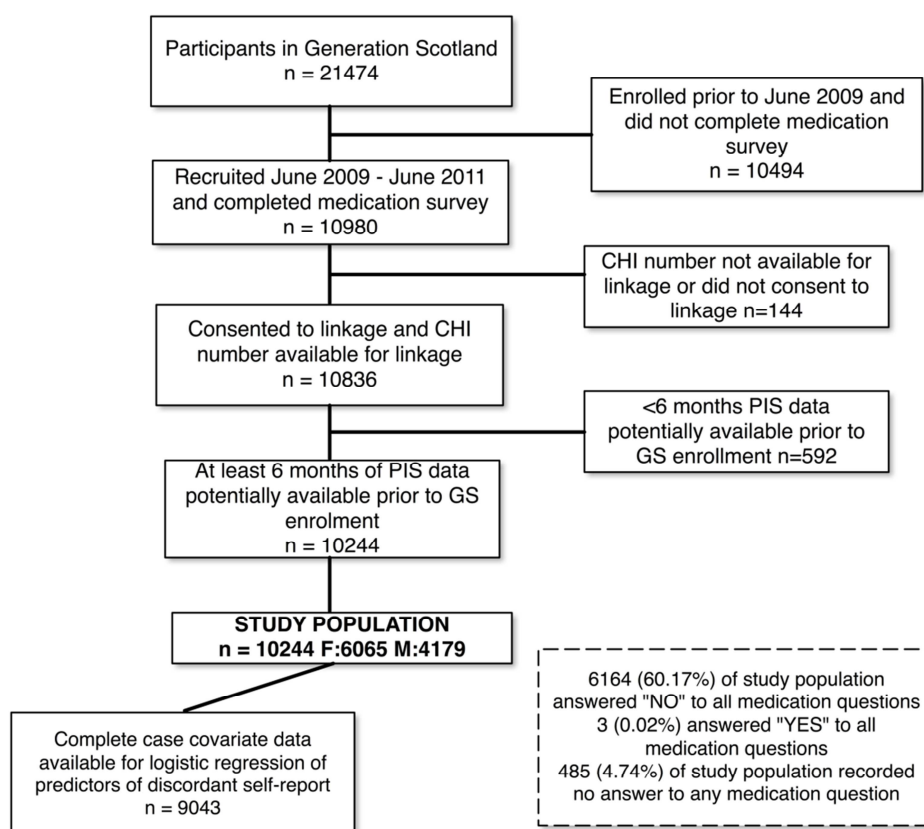
*** Note that a narrower definition of mood stabilizer than that shown in table, which comprised only lithium, sodium valproate, lamotrigine and carbamazepine, returned an agreement of $\kappa=0.29(0.20-0.38)$ at six month time window with sensitivity of 0.21(0.22-0.43)

Table 3. Odds Ratios (with 95% Confidence Intervals) For Factors Associated With Failure To Self-Report Medication Use (False Negatives) As Determined By Prescribing Data As Gold Standard

	Antidepressants	Mood stabilizers	Cholesterol lowering medication	Antihypertensives	Aspirin	Oral contraceptives (females only)
Female sex	0.67 (0.42-1.09)	0.75 (0.24-2.33)	1.62 (0.80-3.30)	1.75 (1.16-2.62)	1.14 (0.52-2.48)	-
Age	0.97 (0.95-0.99)	0.96 (0.91-1.02)	0.95 (0.92-0.99)	0.94(0.92-0.96)	<i>0.94 (0.90-0.99)</i>	1.01 (0.98-1.04)
Affective disorder	<i>0.55 (0.35-0.87)</i>	0.09 (0.02-0.35)	0.72 (0.22-2.42)	0.82 (0.47-1.44)	0.70 (0.19-2.51)	1.31 (0.69-2.49)
Diabetes	-	-	0.42 (0.13-1.40)	0.30 (0.13-0.70)	-	-
Hypertension	-	-	<i>0.28 (0.11-0.71)</i>	0.04 (0.02-0.06)	0.49 (0.23-1.06)	-
Heart disease	-	-	0.30 (0.07-1.25)	0.82 (0.45-1.50)	<i>0.15 (0.03-0.65)</i>	-
No school certificate	0.60 (0.26-1.32)	17.0 (2.3-125.84)	0.45 (0.12-1.72)	0.66 (0.37-1.17)	0.88 (0.28-2.82)	0.65 (0.07-5.89)
Higher education	1.17 (0.70-2.00)	1.27 (0.25-6.35)	1.63 (0.65-1.09)	0.85 (0.54-1.34)	1.27 (0.44-3.64)	1.41 (0.80-2.49)
Smoker	0.90 (0.52-1.54)	0.12 (0.02-0.082)	1.30 (0.45-3.76)	<i>1.84 (1.09-3.11)</i>	1.58 (0.59-4.21)	<i>1.98 (1.13-3.46)</i>
Ex-Smoker	0.66 (0.38-1.11)	0.44 (0.10-2.00)	1.32 (0.59-2.92)	1.40 (0.93-2.12)	0.71 (0.28-1.81)	1.18 (0.65-2.14)
Relationship status - couple	0.89 (0.56-1.41)	2.03(0.59-7.01)	1.31 (0.58-2.97)	0.96 (0.63-1.47)	0.91 (0.40-2.08)	0.78 (0.48-1.28)
General intelligence (g)	0.85 (0.70-1.04)	0.76 (0.46-1.26)	0.85 (0.65-1.11)	1.02 (0.85-1.21)	1.17 (0.83-1.66)	0.92 (0.74-1.15)
Psychological distress (GHQ Likert)	<i>0.98 (0.96-1.00)</i>	0.96 (0.91-1.01)	0.99 (0.95-1.04)	0.99 (0.97-1.01)	1.00 (0.95-1.04)	1.02 (0.99-1.04)

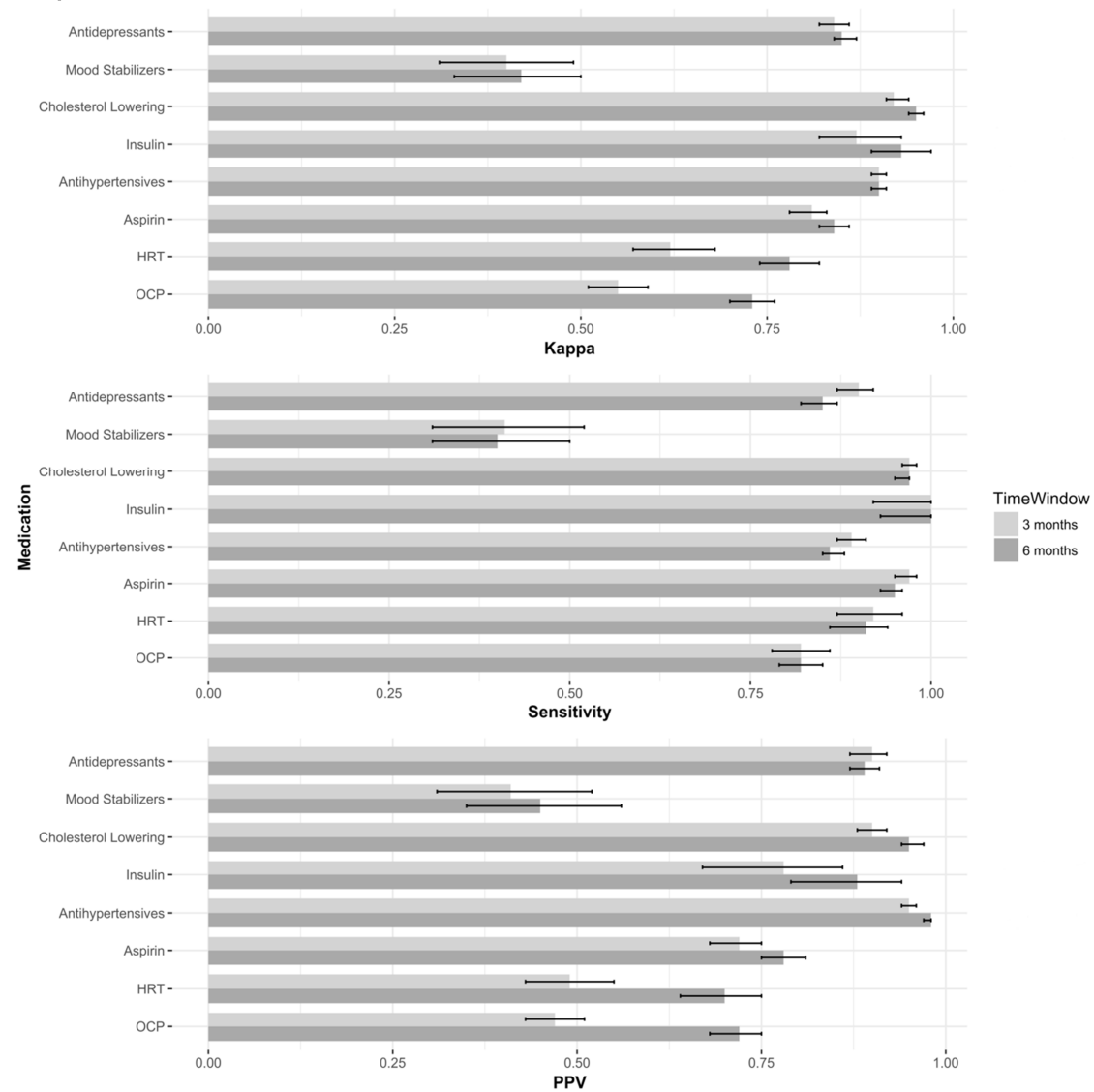
Significant associations are shown in **bold** (alpha=0.05 and adjusted for multiple testing by False Discovery Rate method) and near-significant associations (alpha >0.10) are shown in *italics*. The following factors were used as controls and do not appear in the table : male sex; age 18-39; secondary school education only; no affective disorder found on SCID; no history of self-reported high blood pressure/heart disease/diabetes; smoking status –never smoked; relationship status –single. Insulin and hormone replacement therapy (HRT) are not shown in the table as no significant associations with predictors were found.

Figure 1. Flowchart of derivation of study population, and subset used in logistic regression analysis, from the Generation Scotland cohort.



Abbreviations : GS = Generation Scotland; PIS = Prescribing Information System; CHI = Community Health Index

Figure 2. Agreement and Validity of Medication Self-Report Compared With Prescribing Data As Gold Standard Using Three And Six Month Fixed Time Windows, With 95% Confidence Intervals



Abbreviations : PPV = Positive Predictive Value; HRT = Hormone Replacement Therapy; OCP = Oral Contraceptive Pill.